

JSAI: Designing a Sound, Configurable, and Efficient Static Analyzer for JavaScript

Vineeth Kashyap[†] Kyle Dewey[†] Ethan A. Kuefner[†] John Wagner[‡]
 Kevin Gibbons[‡] John Sarracino[⧻] Ben Wiedermann[⧻] Ben Harkdekopf[†]

University of California Santa Barbara[†]
 {vineeth, kyledewey, eakuefner, benh}@cs.ucsb.edu

University of California Santa Barbara[‡]
 {john_wagner, kgibbons}@umail.ucsb.edu

Harvey Mudd College[⧻]
 {jsarracino@g, benw@cs}.hmc.edu

Abstract

We describe JSAI, an abstract interpreter for JavaScript. JSAI uses novel abstract domains to compute a reduced product of type inference, pointer analysis, string analysis, integer and boolean constant propagation, and control-flow analysis. In addition, JSAI allows for analysis control-flow sensitivity (i.e., context-, path-, and heap-sensitivity) to be modularly configured without requiring any changes to the analysis implementation. JSAI is designed to be provably sound with respect to a specific concrete semantics for JavaScript, which has been extensively tested against existing production-quality JavaScript implementations.

We provide a comprehensive evaluation of JSAI's performance and precision using an extensive benchmark suite. This benchmark suite includes real-world JavaScript applications, machine-generated JavaScript code via Emscripten, and browser addons. We use JSAI's configurability to evaluate a large number of analysis sensitivities (some well-known, some novel) and observe some surprising results. We believe that JSAI's configurability and its formal specifications position it as a useful research platform to experiment on novel sensitivities, abstract domains, and client analyses for JavaScript.

1. Introduction

JavaScript is pervasive. While it began as a client-side webpage scripting language, JavaScript is now used for a wide variety of purposes—for example, to extend the functionality of web browsers in the form of browser addons, to develop desktop applications (e.g., for Windows 8 [6]) and server-side applications (e.g., using Node.js [13]), and to develop mobile phone applications (e.g., for Firefox OS [7]). A growing number of languages, from C to Haskell, can now be compiled to JavaScript [10]. JavaScript's growing prominence means that secure, correct, maintainable and fast JavaScript code is becoming ever more critical. Static analysis traditionally plays a large role in providing these characteristics: it can be used for security auditing, error-checking, debugging, optimization, and program refactoring, among other uses. Thus, a sound, precise static analysis platform for JavaScript can be of enormous advantage.

JavaScript is an inherently dynamic language: it is dynamically typed, object properties (the JavaScript name for object members) can be dynamically inserted and deleted, prototype-based inheri-

tance allows inheritance relations to be changed dynamically, implicit type conversions are abundant and can trigger user-defined code, and more. This dynamism makes static analysis of JavaScript a significant challenge. Compounding this difficulty is the fact that JavaScript analysis is a relatively new endeavor—we as a community have barely begun to explore the many possible approximations and abstractions that balance precision and performance, in order to determine which ones are most appropriate for JavaScript.

The current state-of-the-art static analyses for JavaScript usually take one of two approaches: either (1) an unsound¹ dataflow analysis-based approach using baked-in data abstractions and baked-in context- and heap-sensitivities [18, 26, 32], or (2) a formally-specified type system, proven sound with respect to a specific JavaScript formal semantics but restricted to a small subset of the full JavaScript language [20, 28, 30, 47].

In this work we introduce JSAI, the JavaScript Abstract Interpreter. Our goal is to push the state of the art in JavaScript static analysis along several dimensions. Specifically, our design goals for JSAI are:

Soundness. Our research question is how far we can push sound analysis for full JavaScript while remaining practical, in contrast with most existing full JavaScript analyses which give up on soundness due to JavaScript's complexity. JSAI is based on the theory of abstract interpretation [21], which formally relates the soundness and precision of an abstract semantics (i.e., the static analysis) with a given concrete semantics. Existing proposed concrete semantics for JavaScript turn out to be inadequate for this purpose; we have designed both concrete and abstract semantics for JavaScript specifically for abstract interpretation. We have designed JSAI so that its implementation closely corresponds to its formal specification—it is, in effect, an executable semantics. JSAI handles JavaScript as specified by the ECMA 3 standard [23] (sans `eval` and family; this is further discussed in Section 3.4), along with various language extensions such as Typed Arrays [15].

¹ Most examples of this approach are intentionally unsound as a design decision, in order to handle the difficulties raised by JavaScript analysis. While unsound analysis can be useful for certain purposes, for other purposes (e.g., security auditing of critical code such as browser addons) sound analysis is a definite plus.

Configurability. In order to explore the space of possible approximations and abstractions for JavaScript analysis, we have designed JSAl to be easily configurable in several ways. First, we enable context-, path- and heap-sensitivity of the analysis to be modularly configured without requiring changes to the rest of the analysis implementation, thus making sensitivity an independent concern. None of the existing static analyses for JavaScript have this capability. Doing so requires novel theoretical insights as detailed in the work by Hardekopf et. al. [29]; we provide the first implementation of the insights contained in that paper for a real-world (i.e., non-toy) language. Analysis designers can specify known or novel sensitivities (e.g., k -CFA, object sensitivity, property simulation) by implementing a simple API that can then be plugged into the analysis. We have implemented over a dozen sensitivities in this manner, each of them requiring only 5–20 lines of code. Secondly, the string and number abstract domains used by the analysis are designed to be easily swapped out for new, experimental abstract domains. Strings and numbers are prevalent in JavaScript, and therefore designing the right abstractions for these can have a useful impact on analysis precision and performance. We have designed these domains each implement a specific API which is used by the rest of the analysis; any abstract domain implementing these APIs can be used in their place.

Efficiency. JSAl is designed to be competitive with existing JavaScript analyses in terms of performance while still meeting its goals of soundness and configurability. It incorporates various analysis optimizations to enable it to scale to real-world JavaScript programs of non-trivial size. JSAl is comparable in performance to TAJs [33, 34], the most closely related JavaScript analyzer, while being significantly more configurable and being based on a formalized semantics.

The contributions of the JSAl project include complete formalisms for a concrete and abstract semantics for JavaScript along with implementations of concrete and abstract interpreters based on these semantics. While concrete semantics for JavaScript have been proposed before, ours is the first designed specifically for abstract interpretation. Our abstract semantics is the first formal abstract semantics for JavaScript in the literature. The abstract interpreter implementation is the first available static analyzer for JavaScript that provides easy configurability as a design goal. All these contributions are available freely for download as supplementary materials². Thus JSAl provides a research platform to experiment with a variety of context-, path- and heap-sensitivities and abstract domains, and it provides a solid foundation on which to build multiple client analyses for JavaScript. In fact, JSAl has been used to build a security auditing tool for browser addons [36], and to experiment with type refinement as a strategy to improve analysis precision [37]. The contributions of this paper include:

- The design of a JavaScript intermediate language and concrete semantics intended specifically for abstract interpretation (Section 3.1).
- The design of an abstract semantics that enables configurable, sound abstract interpretation for JavaScript (Section 3.2). This abstract semantics represents a reduced product [22] of type inference, pointer analysis, string analysis, integer and boolean constant propagation, and control-flow analysis—all working together in carefully-designed harmony to enable precise tracking of data- and control-flow within a JavaScript program.
- Novel abstract string and object domains for JavaScript analysis (Section 3.3).
- Two novel context sensitivities for JavaScript (Section 4).

- An evaluation of JSAl’s performance and precision on the most comprehensive suite of benchmarks for JavaScript static analysis that we are aware of, including browser addons, machine-generated programs via Emscripten [5], and open-source JavaScript programs (Section 5). We showcase JSAl’s configurability by evaluating a large number of context- and heap-sensitivities, and point out novel insights from the results.

We preface these contributions with a discussion of related work (Section 2) and conclude with plans for future work (Section 6).

2. Related Work

In this section we discuss existing static and hybrid approaches to analyzing JavaScript, and also discuss previous efforts to formalize JavaScript semantics. Finally, we discuss previous efforts for configurable static analysis.

2.1 JavaScript Analyses

Previous work on analyzing JavaScript programs either gives up soundness, or analyzes a restricted subset of the language, or both. None of the previous JavaScript analyses target configurability.

Various previous work [16, 24, 25, 32, 39, 46, 47] proposes different subsets of the JavaScript language, and provides analyses for that subset. These analyses range from type inference, to pointer analysis, to numeric range and kind analysis for program optimization. None of these handle the full complexities of JavaScript.

Unsound analysis can be useful under certain circumstances, and there have been intentionally unsound analyses [4, 18, 40] proposed for JavaScript. Other works [26, 32] take a best-effort approach to soundness, without any assurance that the analysis is actually sound.

Several type systems [20, 28, 30, 47] have been proposed to retrofit JavaScript (or subsets thereof) with static types. Guha et. al. [28] propose a novel combination of type systems and flow analysis, and apply it to JavaScript. Chugh et. al. [20] propose a flow-sensitive refinement type system designed to allow typing of common JavaScript idioms. These type systems require programmer annotations and cannot be used as-is on real-world JavaScript programs, as opposed to our fully automatic approach.

Combinations of static analysis with dynamic checks [19, 25] have been proposed to handle JavaScript—these systems statically analyze a subset of JavaScript under certain assumptions and use runtime checks to enforce these assumptions. Schäfer et al. [43] use a dynamic analysis to determine information that can be leveraged to scale static analysis for JavaScript. These ideas can usefully supplement our static techniques.

Jensen et. al. [33] present a state-of-the-art static analysis for JavaScript. They have since improved on this analysis in several ways [34, 35]; we refer to this entire body of work as TAJs. TAJs translates JavaScript programs into a flowgraph-based IR upon which the analysis is run. While TAJs is intended to be sound, there is no attempt to formalize the translation to the IR, the semantics of the IR, or the analysis itself. In fact, while formalizing our work we found some subtle and previously unknown soundness bugs in TAJs. TAJs also does not have the design goal of configurability, therefore it does not allow for tunable control-flow sensitivity or modularly replacing various abstract domains.

2.2 JavaScript formalisms

None of the previous work on static analysis of JavaScript has formally specified the analysis or attempted to prove soundness. However, there has been previous work on providing JavaScript with a formal semantics.

Maffei et. al [41] give a structural smallstep operational semantics directly to the full JavaScript language (except a few

² At this URL: <http://cs.ucsb.edu/~vineeth/axiv/jsai.zip>

constructs). Lee et. al [38] propose SAFE, a semantic framework that provides structural bigstep operational semantics to JavaScript, based directly on the ECMAScript specification. Due to their size and complexity, neither of these semantic formulations are suitable for direct translation into an abstract interpreter.

Guha et. al [27] propose a core calculus approach to provide semantics to JavaScript—they provide a desugarer (parts of which are formally specified) from JavaScript to a core calculus called λ_{JS} , which has a smallstep structural operational semantics. Their intention was to provide a minimal core calculus that would ease proving soundness for type systems, thus placing all the complexity in the desugarer. However, their core calculus is too low-level to perform a precise and scalable static analysis (for example, some of the semantic structure that is critical for a precise analysis is lost, and their desugaring causes a large code bloat). We also use the core calculus approach; however, our own intermediate language notJS is designed to be in a sweet-spot—the complexity is shared between the translator and the notJS semantics with the emphasis placed on static analysis. In addition, we use an abstract machine-based semantics rather than a structural semantics, which (as described later) is what enables configurable analysis sensitivity.

2.3 Configurable Analysis

Sergey et al. [44] describe monad-based techniques for abstracting certain characteristics of an abstract interpreter, allowing the analysis behavior to be configured by plugging in different independently-specified monads. They demonstrate their technique for lambda calculus and for Featherweight Java. However, their work does not allow analysis sensitivity to be configured in this way (and, in fact, their described analyses have intractable complexity). Our work is complementary, in that we show how to make the analysis sensitivity configurable in a manner that allows the analysis tractability to be controlled. In addition, we demonstrate our technique on a complete real-world language, JavaScript.

3. JSAI Design

We break our discussion of the JSAI design into three main components: (1) the design of an intermediate representation for JavaScript programs, called notJS, along with its concrete semantics; (2) the design of an abstract semantics for notJS that yields the reduced product of a number of essential sub-analyses and also enables configurable analysis; and (3) the design of novel abstract domains for JavaScript analysis. We conclude with a discussion of various options for handling dynamic code injection.

The intent of this section is to discuss the design decisions that went into JSAI, rather than giving a comprehensive description of the various formalisms (e.g., the translation from JavaScript to notJS, the concrete semantics of notJS, and the abstract semantics of notJS). All of these formalisms, along with their implementations, appear in the supplementary materials.

3.1 Designing the notJS Intermediate Language

Our soundness goal motivates the use of formal specifications for both concrete JavaScript semantics and our abstract analysis semantics. Our approach is to define an intermediate language called notJS, along with a formally-specified translation from JavaScript to notJS. We then give notJS a formal concrete semantics upon which we base our abstract interpreter.³

Figure 1 shows the abstract syntax of notJS, which was carefully designed with the ultimate goal of making abstract interpretation simple, precise, and efficient. JavaScript’s builtin objects (e.g.,

Math) and methods (e.g., isNaN) are properties of the global object constructed prior to a program’s execution, thus they are not a part of the language syntax.

$$\begin{aligned}
 n &\in \text{Num} \quad b \in \text{Bool} \quad \text{str} \in \text{String} \quad x \in \text{Variable} \quad \ell \in \text{Label} \\
 s &\in \text{Stmt} ::= \vec{s}_i \mid \text{if } e \, s_1 \, s_2 \mid \text{while } e \, s \mid x := e \mid e_1.e_2 := e_3 \\
 &\quad \mid x := e_1(e_2, e_3) \mid x := \text{toobj } e \mid x := \text{del } e_1.e_2 \\
 &\quad \mid x := \text{newfun } m \, n \mid x := \text{new } e_1(e_2) \mid \text{throw } e \\
 &\quad \mid \text{try-catch-fin } s_1 \, x \, s_2 \, s_3 \mid \ell \, s \mid \text{jump } \ell \, e \mid \text{for } x \, e \, s \\
 e &\in \text{Exp} ::= n \mid b \mid \text{str} \mid \text{undef} \mid \text{null} \mid x \mid m \mid e_1 \oplus e_2 \mid \odot e \\
 d &\in \text{Decl} ::= \text{decl } \vec{x}_i = \vec{e}_i \text{ in } s \\
 m &\in \text{Meth} ::= (\text{self}, \text{args}) \Rightarrow d \mid (\text{self}, \text{args}) \Rightarrow s \\
 \oplus &\in \text{BinOp} ::= + \mid - \mid \times \mid \div \mid \% \mid \ll \mid \gg \mid \ggg \mid < \mid \leq \mid \& \\
 &\quad \mid ' \mid \vee \mid \text{and} \mid \text{or} \mid ++ \mid < \mid \leq \mid \approx \mid \equiv \mid . \\
 &\quad \mid \text{instanceof} \mid \text{in} \\
 \odot &\in \text{UnOp} ::= - \mid \sim \mid \neg \mid \text{typeof} \mid \text{isprim} \mid \text{tobool} \\
 &\quad \mid \text{tostr} \mid \text{tonum}
 \end{aligned}$$

Figure 1: The abstract syntax of notJS provides canonical constructs that simplify JavaScript’s behavior. The vector notation represents (by abuse of notation) an ordered sequence of unspecified length n , where i ranges from 0 to $n - 1$.

An important design decision we made is to separate the language into pure expressions ($e \in \text{Exp}$) that are guaranteed to terminate without throwing an exception, and impure statements ($s \in \text{Stmt}$) that do not have these guarantees. This decision directly impacts the formal semantics and implementation of notJS, a further discussion of which appears later in this section. This is the first IR for JavaScript we are aware of that makes this design choice—it is a more radical choice than might first be apparent, because JavaScript’s implicit conversions make it difficult to enforce this separation. The IR was carefully designed to make this possible. Some other design decisions of note include making JavaScript’s implicit conversions (which are complex and difficult to reason about, involving multiple steps and alternatives depending on the current state of the program) explicit in notJS; leaving certain JavaScript constructs unlowered to allow for a more precise abstract semantics (e.g., the `for...in` loop, which we leave mostly intact as `for x e s`); and simplifying method calls to make the implicit `this` parameter and `arguments` object explicit—this is often, but not always, the address of a method’s receiver object, and its value can be non-intuitive, while `arguments` provides a form of reflection providing access to a method’s arguments.

Given the notJS abstract syntax, we need to design a formal concrete semantics that (together with the translation to notJS) captures JavaScript behavior. We have two main criteria: (1) the semantics should be specified in a manner that can be directly converted into an implementation, allowing us to test its behavior against actual JavaScript implementations; (2) looking ahead to the abstract version of the semantics (which defines our analysis), the semantics should be specified in a manner that allows for configurable sensitivity. These requirements lead us to specify the notJS semantics as an abstract machine-based smallstep operational semantics. One can think of this semantics as an infinite state transition system, wherein we formally define a notion of *state* and a set of *transition rules* that connect states. The semantics is implemented by turning the state definition into a data structure (e.g., a Scala class) and the transition rules into functions that transform a given state into the next state. The concrete interpreter starts with an initial state (containing the start of the program and all of the builtin JavaScript methods and objects), and continually computes the next state until the program finishes.

³Guha et al [27] use a similar approach, but our IR design and formal semantics are different. See Section 2 for a discussion of the differences between our two approaches.

Further Design Discussion. Previous efforts to give JavaScript a formal concrete semantics all use either bigstep or smallstep structural operational semantics. However, a smallstep abstract machine semantics is more suited for abstract interpretation (particularly to enable configurability in the form of tunable control-flow sensitivity and straightforward implementation). Our semantics is actually not completely smallstep: expressions are evaluated in a big-step style, which means they are evaluated via a recursive traversal of their abstract syntax tree (AST), similar to most AST-based interpreters. This is made possible by our separation of expressions (pure, terminating) from statements (impure, potentially non-terminating). While this separation might be standard for simpler languages, it took careful design of the notJS IR to enable this separation for JavaScript.

Initially we designed notJS so that there was no separation between statements and expressions, and side-effects and exceptions could happen anywhere. We designed the corresponding semantics to be in completely smallstep style. As opposed to our current design (which keeps expressions separate and guarantees they are pure), the initial design had three times as many semantic continuations, and more complicated reasoning for the semantic rules.

We omit further details of the concrete semantics both for space and because they are almost redundant with the abstract semantics described in the next section. The main difference between the two is that the abstract state employs sets in places where the concrete state employs singletons, and the abstract transition rules are nondeterministic whereas the concrete rules are deterministic. Both of these differences are because the abstract semantics over-approximates the concrete semantics.

Testing the Semantics. We tested the translation, semantics, and implementation thereof by comparing its behavior with that of an actual JavaScript engine, SpiderMonkey [14]. We constructed a test suite of over a million JavaScript programs, most of which were randomly generated. However, 243 of the programs in the test suite were either hand-crafted to exercise various parts of the semantics, or taken from existing JavaScript programs used to test commercial JavaScript implementations. We then ran all of the tests on SpiderMonkey and on our concrete interpreter, and we verified that they produce identical output. While we can never completely guarantee that the notJS semantics matches the ECMA specification, we can do as well as any JavaScript implementation, which goes through the same sort of testing process.

3.2 Designing the Abstract Semantics

The JavaScript static analysis is defined as an abstract semantics for notJS that over-approximates the notJS concrete semantics. The analysis is implemented by computing the set of all abstract states reachable from a given initial state by following the abstract transition rules. The analysis contains some special machinery that provides configurable sensitivity. We illustrate our approach via a worklist algorithm that ties these concepts together:

The static analysis performed by this worklist algorithm is determined by the definitions of the abstract semantic states $\hat{\zeta} \in State^\#$, the abstract transition rules $next_states(\hat{\zeta}) \in State^\# \rightarrow \mathcal{P}(State^\#)$, and the knob that configures the analysis sensitivity $trace(\hat{\zeta})$. We discuss each of these aspects in turn.

Abstract Semantic Domains. Figure 2 shows our definition of an abstract state for notJS. An abstract state $\hat{\zeta}$ consists of a *term* that is either a notJS statement or an abstract value that is the result of evaluating a statement; an *environment* that maps variables to (sets of) addresses; a *store* mapping addresses to either abstract values, abstract objects, or sets of continuations (to enforce computability for abstract semantics that use semantic continuations, as per Van Horn and Might [48]); and finally a *continuation stack* that

Algorithm 1 The JSAI worklist algorithm

```

1: put the initial abstract state  $\hat{\zeta}_0$  on the worklist
2: initialize map  $partition : Trace \rightarrow State^\#$  to empty
3: repeat
4:   remove an abstract state  $\hat{\zeta}$  from the worklist
5:   for all abstract states  $\hat{\zeta}'$  in  $next\_states(\hat{\zeta})$  do
6:     if  $partition$  does not contain  $trace(\hat{\zeta}')$  then
7:        $partition(trace(\hat{\zeta}')) = \hat{\zeta}'$ 
8:       put  $\hat{\zeta}'$  on worklist
9:     else
10:       $\hat{\zeta}_{old} = partition(trace(\hat{\zeta}'))$ 
11:       $\hat{\zeta}_{new} = \hat{\zeta}_{old} \sqcup \hat{\zeta}'$ 
12:      if  $\hat{\zeta}_{new} \neq \hat{\zeta}_{old}$  then
13:         $partition(trace(\hat{\zeta}_{new})) = \hat{\zeta}_{new}$ 
14:        put  $\hat{\zeta}_{new}$  on worklist
15:      end if
16:    end if
17:  end for
18: until worklist is empty

```

$$\begin{aligned}
\hat{n} &\in Num^\# & \hat{str} &\in String^\# & \hat{a} &\in Address^\# & \hat{o} &\in UnOp^\# & \hat{\oplus} &\in BinOp^\# \\
\hat{\zeta} \in State^\# &= Term^\# \times Env^\# \times Store^\# \times Kont^\# \\
\hat{t} \in Term^\# &= Decl + Stmt + Value^\# \\
\hat{\rho} \in Env^\# &= Variable \rightarrow \mathcal{P}(Address^\#) \\
\hat{\sigma} \in Store^\# &= Address^\# \rightarrow (BValue^\# + Object^\# + \mathcal{P}(Kont^\#)) \\
\hat{bv} \in BValue^\# &= Num^\# \times \mathcal{P}(Bool) \times String^\# \times \mathcal{P}(Address^\#) \times \\
&\quad \mathcal{P}(\{\text{null}\}) \times \mathcal{P}(\{\text{undef}\}) \\
\hat{o} \in Object^\# &= (String^\# \rightarrow BValue^\#) \times \mathcal{P}(String) \times \\
&\quad (String \rightarrow (BValue^\# + Class + \mathcal{P}(Closure^\#))) \\
c \in Class &= \{\text{function, array, string, boolean, number, date,} \\
&\quad \text{error, regexp, arguments, object, } \dots\} \\
\hat{clo} \in Closure^\# &= Env^\# \times Meth \\
\hat{ev} \in EValue^\# &::= \text{exc } bv \\
\hat{fv} \in JValue^\# &::= \text{jmp } \ell \ \hat{bv} \\
\hat{v} \in Value^\# &= BValue^\# + EValue^\# + JValue^\# \\
\hat{\kappa} \in Kont^\# &::= \widehat{\text{haltK}} \mid \widehat{\text{seqK}} \ \vec{s_i} \ \hat{\kappa} \mid \widehat{\text{whileK}} \ e \ s \ \hat{\kappa} \mid \widehat{\text{lblK}} \ \ell \ \hat{\kappa} \\
&\quad \mid \widehat{\text{forK}} \ \vec{str_i} \ x \ s \ \hat{\kappa} \mid \widehat{\text{retK}} \ x \ \hat{\rho} \ \hat{\kappa} \ \text{ctor} \mid \widehat{\text{retK}} \ x \ \hat{\rho} \ \hat{\kappa} \ \text{call} \\
&\quad \mid \widehat{\text{tryK}} \ x \ s \ s \ \hat{\kappa} \mid \widehat{\text{catchK}} \ s \ \hat{\kappa} \mid \widehat{\text{finK}} \ \vec{v} \ \hat{\kappa} \mid \widehat{\text{addrK}} \ \hat{a}
\end{aligned}$$

Figure 2: Abstract semantic domains for notJS.

represents the remaining computations to perform—one can think of this component as analogous to a runtime stack that remembers computations that should be completed once the current computation is finished.

Abstract values are either exception/jump values ($EValue^\#$, $JValue^\#$), used to handle non-local control-flow, or base values ($BValue^\#$), used to represent JavaScript values. Base values are a tuple of abstract numbers, booleans, strings, addresses, null, and undefined; each of these components is a lattice. Base values are defined as tuples because the analysis over-approximates the concrete semantics, and thus cannot constrain values to be only a single type at a time. These value tuples yield a type inference analysis: any component of this tuple that is a lattice \perp represents a type that

this value cannot contain. Base values do not include function closures, because functions in JavaScript are actually objects. Instead, we define a class of abstract objects that correspond to functions and that contain a set of closures that are used when that object is called as a function. We describe our novel abstract object domain in more detail in Section 3.3.

Each component of the tuple also represents an individual analysis: the abstract number domain determines a number analysis, the abstract string domain determines a string analysis, the abstract addresses domain determines a pointer analysis, etc. Composing the individual analyses represented by the components of the value tuple is not a trivial task; a simple cartesian product of these domains (which corresponds to running each analysis independently, without using information from the other analyses) would be imprecise to the point of being useless. Instead, we specify a reduced product [22] of the individual analyses, which means that we define the semantics so that each individual domain can take advantage of the other domains' information to improve their results. The abstract number and string domains are intentionally unspecified in the semantics; they are configurable. We discuss our specific implementations of the abstract string domain in Section 3.3.

Together, all of these abstract domains define a set of simultaneous analyses: control-flow analysis (for each call-site, which methods may be called), pointer analysis (for each object reference, which objects may be accessed), type inference (for each value, can it be a number, a boolean, a string, **null**, **undef**, or a particular class of object), and extended versions of boolean, number, and string constant propagation (for each boolean, number and string value, is it a known constant value). These analyses combine to give detailed control- and data-flow information forming a fundamental analysis that can be used by many possible clients (e.g., error detection, program slicing, secure information flow).

Abstract Transition Rules. Figure 3 describes a small subset of the abstract transition rules, to give their flavor. To compute $\text{next.states}(\xi)$, the components of ξ are matched against the premises of the rules to find which rule(s) are relevant; that rule then describes the next state (if multiple rules apply, then there will be multiple next states). The rules 1, 2, and 3 deal with sequences of statements. Rule 1 says that if the state's term is a sequence, then pick the first statement in the sequence to be the next state's term; then take the rest of the sequence and put it in a **seqK** continuation for the next state, pushing it on top of the continuation stack. Rule 2 says that if the state's term is a base value (and hence we have completed the evaluation of a statement), take the next statement from the **seqK** continuation and make it the term for the next state. Rule 3 says that if there are no more statements in the sequence, pop the **seqK** continuation off of the continuation stack. The rules 4 and 5 deal with conditionals. Rule 4 says that if the guard expression evaluates to an abstract value that over-approximates **true**, make the **true** branch statement the term for the next state; rule 5 is similar except it takes the **false** branch. Note that these rules are nondeterministic, in that the same state can match both rules.

Configurable Sensitivity. To enable configurable sensitivity, we build on the insights of Hardekopf et al [29]. We extend the abstract state to include an additional component from a *Trace* abstract domain. The worklist algorithm uses the *trace* function to map each abstract state to its trace, and joins together all reachable abstract states that map to the same trace (see lines 10–11 of Algorithm 1). The definition of *Trace* is left to the analysis designer; different definitions yield different sensitivities. For example, suppose *Trace* is defined as the set of program points, and an individual state's trace is the current program point. Then our worklist algorithm computes a flow-sensitive, context-insensitive analysis: all states at the same program point are joined together, yielding

	Current State ξ	Next State ξ'
1	$\langle s :: \vec{s}_i, \hat{\rho}, \hat{\sigma}, \hat{\kappa} \rangle$	$\langle s, \hat{\rho}, \hat{\sigma}, \widehat{\text{seqK}} \vec{s}_i \hat{\kappa} \rangle$
2	$\langle \widehat{bv}, \hat{\rho}, \hat{\sigma}, \widehat{\text{seqK}} s :: \vec{s}_i \hat{\kappa} \rangle$	$\langle s, \hat{\rho}, \hat{\sigma}, \widehat{\text{seqK}} \vec{s}_i \hat{\kappa} \rangle$
3	$\langle \widehat{bv}, \hat{\rho}, \hat{\sigma}, \widehat{\text{seqK}} \epsilon \hat{\kappa} \rangle$	$\langle \widehat{bv}, \hat{\rho}, \hat{\sigma}, \hat{\kappa} \rangle$
4	$\langle \text{if } e \ s_1 \ s_2, \hat{\rho}, \hat{\sigma}, \hat{\kappa} \rangle$	$\langle s_1, \hat{\rho}, \hat{\sigma}, \hat{\kappa} \rangle \quad \text{if } \text{true} \in \pi_{\hat{b}}(\llbracket e \rrbracket)$
5	$\langle \text{if } e \ s_1 \ s_2, \hat{\rho}, \hat{\sigma}, \hat{\kappa} \rangle$	$\langle s_2, \hat{\rho}, \hat{\sigma}, \hat{\kappa} \rangle \quad \text{if } \text{false} \in \pi_{\hat{b}}(\llbracket e \rrbracket)$

Figure 3: A small subset of the abstract semantics rules for JSAI. Each smallstep rule describes a transition relation from one abstract state ξ to the next state ξ' . The phrase $\pi_{\hat{b}}(\llbracket e \rrbracket)$ means to evaluate expression e to an abstract base value, then project out its boolean component.

one state per program point. Suppose we redefine *Trace* to be sequences of program points, and an individual state's trace to be the last k call-sites. Then our worklist algorithm computes a flow-sensitive, k -CFA context-sensitive analysis. Arbitrary sensitivities can be defined in this manner solely by redefining *Trace*, without affecting the worklist algorithm or the abstract transition rules. We explore a number of possibilities in Section 5.

While most static analyses are built on top of the control-flow graph (CFG), JSAI instead employs semantic continuations inside the abstract states. One reason is so we can employ the techniques of [29] to enable configurable sensitivity. These techniques are also an important reason that we used abstract machine-based semantics instead of structural semantics as in all previous formalizations of JavaScript—a structural semantics would not allow us to configure the sensitivity in this way. Another reason for using semantic continuations is that JavaScript contains a great deal of indirect and non-local control-flow (due to higher-order functions, implicit exceptions, etc), thus much of the control-flow is non-obvious and basing the analysis on a CFG would require a great deal of ad-hoc patches during the analysis. One class of soundness bugs relating to **try-catch-finally** that we found in TAJs, which does use a CFG, was due to exactly this issue.

3.3 Novel Abstract Domains

JSAI allows configurable abstract number and string domains, but we also provide default domains based on our experience with JavaScript analysis. We motivate and describe our default abstract string domain here. We also describe our novel abstract object domain, which is an integral part of the JSAI abstract semantics.

Abstract Strings. Our initial abstract string domain *String*[#] was a simple string constant domain—the elements were either constant strings or \top , representing an unknown string. We extended that domain to separate unknown strings into two categories: strings that are definitely numbers, and strings that are definitely not numbers (borrowing from TAJs [33]). This separation helps when analyzing arrays: arrays are just objects where array indices are represented with numeric string properties such as "0", "1", etc, but they also have non-numeric properties like "length". However, this initial string domain was inadequate.

In particular, we discovered a need to express that a string is *not* contained within a given hard-coded set of strings. Consider the property lookup $x := \text{obj}[y]$, where y is a variable that resolves to an unknown string. Because the string is unknown, the analysis is forced to assign to x not only the lattice join of all values contained in obj , but also the lattice join of all the values contained in all prototypes of obj , due to the rules of prototype-based inheritance. Almost all object prototype chains terminate in one of the builtin objects contained in the global object (`Object.prototype`, `Array.prototype`, etc); these builtin objects contain the builtin values and methods. Thus, all of these builtin values and methods are returned for any object property access based on an unknown string,

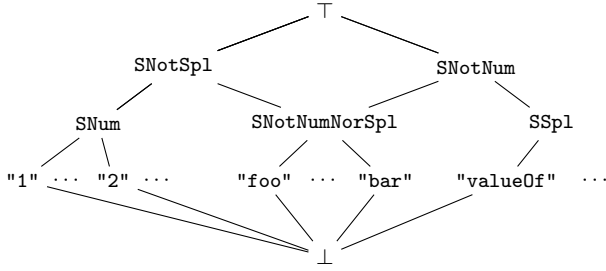


Figure 4: Our default string abstract domain, further explained in Section 3.3.

severely polluting the results. One possible way to mitigate this problem is to use an expensive domain that can express arbitrary complements (i.e., express that a string is *not* contained in some arbitrary set of strings). Instead, we extend the string domain to separate out *special* strings (`valueOf`, `toString` etc.) from the rest; these special strings are drawn from property names of builtin values and methods. We can thus express that a string has an unknown value that is *not* one of the special values. This is a practical solution that improves precision at minimal cost.

The new abstract string domain depicted in Figure 4 (that separates unknown strings into numeric, non-numeric and special strings) was simple to implement due to JSAI’s configurable architecture; it did not require changes to any other parts of the implementation despite the pervasive use of strings in all aspects of JavaScript semantics.

Abstract Objects. We highlight the abstract domain *Object*[#] of Figure 2 as a novel contribution. Previous JavaScript analyses model abstract objects as a tuple containing (1) a map from property names to values; and (2) a list of definitely present properties (necessary because property names are just strings, and objects can be modified using unknown strings as property names). However, according to the ECMA standard objects can be of different *classes*, such as functions, arrays, dates, regexps, etc. While these are all objects and share many similarities, there are semantic differences between objects of different classes. For example, the `length` property of array objects has semantic significance—assigning a value to `length` can implicitly add or delete properties to the array object, and certain values cannot be assigned to `length` without raising a runtime exception. Non-array objects can also have a `length` field, but assigning to that field will have no other effect. The object’s class dictates the semantics of property enumeration, update and delete operations on an object. Thus, the analysis must track what classes an abstract object may belong to in order to accurately model these semantic differences. If abstract objects can belong to arbitrary sets of classes, this tracking and modeling becomes extremely complex, error-prone, and inefficient.

Our innovation is to add a map as the third component of abstract objects that contains class-specific values. This component also records which class an abstract object belongs to. Finally, the semantics is designed so that any given abstract object must belong to exactly one class. This is enforced by assigning abstract addresses to objects based not just on their static allocation site and context, but also on the constructor used to create the object (which determines its class). The resulting abstract semantics is much simpler, more efficient, and precise.

3.4 Handling `eval` and Similar Constructs

Dynamically injected code is the bane of static analysis. JavaScript contains `eval`, which executes an arbitrary string as code.⁴ The notJS IR does not contain an explicit `eval` instruction because `eval` is a builtin method of the global object, rather than being a JavaScript instruction.

There are several possible strategies to handle `eval` for static analysis. For example, we could *disallow* the use of `eval` altogether. In some application domains this is a legitimate strategy—e.g., browser addons must pass through a vetting process to be added to official repositories, and this process strongly discourages `eval`; also, machine-generated JavaScript a la Emscripten [5] rarely contains `eval`. There are also methods to automatically [35] or semi-automatically [42] *eliminate* `eval` in most real-world scenarios. Alternatively, the analysis can make assumptions about the runtime behavior of the `eval` statement, and the program or runtime can be modified to *check* or *enforce* these assumptions, e.g., by running `eval` inside a sandbox. Such runtime checks are used by the staged analysis proposed by Chugh et al. [19]. Finally, the static analysis can initially ignore dynamic code injection, and the runtime can be modified to have the analysis *patch* itself to soundly handle the newly-available information; this is similar to the strategy proposed for handling Java analysis in the presence of dynamic class loading [31].

In this work we do not innovate on methods for handling `eval`; we simply use the *disallow* strategy and have the analysis output a warning if `eval` could potentially have been called. For an important and growing class of JavaScript programs, e.g., browser addons and machine-generated programs, the *disallow* strategy is a sensible choice—none of the 28 real-world benchmarks in our benchmark suite use `eval`. More comprehensive methods for handling `eval` are complementary to JSAI and can be incorporated without significant modifications to the current design.

4. Showcasing JSAI’s Configurability

The primary motivation for making JSAI configurable is to allow analysis designers to explore different possibilities for approximation and abstraction. One important dimension to explore is *context-sensitivity*: how the (potentially infinite) possible method call instances are partitioned and merged into a finite number of abstract instances. The context-sensitivity strategy used by an analysis can greatly influence the analysis precision and performance. The current state of the art for JavaScript static analysis has explored only a few possible context-sensitivity strategies, all of which are baked into the analysis and difficult to change.

We take advantage of JSAI’s configurability to define and evaluate a much larger selection of context-sensitivities than has ever been evaluated before in a single paper. Because of JSAI’s design, specifying each sensitivity takes only 5–20 lines of code, making this process substantially more feasible than existing analysis implementations (where each sensitivity would have to be hard-coded into the analysis from scratch). The analysis designer specifies a sensitivity by instantiating a particular instance of *Trace*; all abstract states with the same trace will be merged together. For context-sensitivity, we define *Trace* to include some notion of the calling context, so that states in the same context are merged while states in different contexts are kept separate.

We implement six main context-sensitivity strategies, each parameterized in various ways, yielding a total of 56 different forms of context-sensitivity. All of our sensitivities are flow-sensitive (JavaScript’s dynamic nature means that flow-insensitive analyses

⁴ There are related constructs with similar functionality; we refer to all of them as `eval` for convenience.

tend to have terrible precision). We empirically evaluate all of these strategies in Section 5; here we define the six main strategies. Four of the six strategies are known in the literature, while two are novel to this paper. The novel strategies are based on two hypotheses about context definitions that might provide a good balance between precision and performance. Our empirical evaluation demonstrates that these hypotheses are false, i.e., they do not provide any substantial benefit. We include them here not as examples of good sensitivities to use, but rather to demonstrate that JSAI makes it easy to formulate and test hypotheses about analysis strategies—each novel strategy took only 15–20 minutes to implement. The strategies we defined are as follows, where the first four are known and the last two are novel:

Context-insensitive. All calls to a given method are merged. We define the context component of *Trace* to be a unit value, so that all contexts are the same.

Stack-CFA. Contexts are distinguished by the list of call-sites on the call-stack. This strategy is k -limited to ensure there are only a finite number of possible contexts. We define the *Trace* component to contain the top k call-sites.

Acyclic-CFA. Contexts are distinguished the same as Stack-CFA, but instead of k -limiting we collapse recursive call cycles. We define *Trace* to contain all call-sites on the call-stack, except that cycles are collapsed.

Object-sensitive. Contexts are distinguished by a list of addresses corresponding to the chain of receiver objects (corresponding to full-object-sensitivity in Smaragdakis et al. [45]). We define *Trace* to contain this information (k -limited to ensure finite contexts).

Signature-CFA. Type information is important for dynamically-typed languages, so intuitively it seems that type information would make good contexts. We hypothesize that defining *Trace* to record the types of a call’s arguments would be a good context-sensitivity, so that all calls with the same types of arguments would be merged.

Mixed-CFA. Object-sensitivity uses the address of the receiver object. However, in JavaScript the receiver object is often the global object created at the beginning of the program execution. Intuitively, it seems this would mean that object sensitivity might merge many calls that should be kept separate. We hypothesized that it might be beneficial to define *Trace* as a modified object-sensitive strategy—when object-sensitivity would use the address of the global object, this strategy uses the current call-site instead.

5. Evaluation

In this section we evaluate JSAI’s precision and performance for a range of context-sensitivities as described in Section 4, for a total of 56 distinct sensitivities. We run each sensitivity on 28 benchmarks collected from four different application domains and analyze the results, yielding surprising observations about context-sensitivity and JavaScript. We also briefly evaluate JSAI as compared to TAJIS [33], the most comparable existing JavaScript analysis.

5.1 Implementation and Methodology

We implement JSAI using Scala version 2.10. We provide a model of the DOM, event handling loop, and other APIs used in our benchmarks. The baseline analysis sensitivity we evaluate is **fs** (flow-sensitive, context-insensitive); all of the other evaluated sensitivities are strictly more precise than **fs**. The other sensitivities are: $k.h$ -**stack**, $k.h$ -**acyclic**, $k.h$ -**obj**, $k.h$ -**sig**, and $k.h$ -**mixed**, where k is the context depth for k -limiting and h is the heap-sensitivity (i.e., the prefix of the context used to distinguish abstract addresses). The

parameters k and h vary from 1 to 5 and $h \leq k$, because the heap sensitivity is always a prefix of the context sensitivity.

We use a comprehensive benchmark suite to evaluate the sensitivities. Most prior work on JavaScript static analysis has been evaluated only on the standard SunSpider [11] and V8 [12] benchmarks, with a few micro-benchmarks thrown in. We evaluate JSAI on these standard benchmarks, but we also include real-world representatives from a diverse set of JavaScript application domains. We choose seven representative programs from each domain, for a total of 28 programs. We partition the programs into four categories, described below. For each category, we provide the mean size of the benchmarks in the suite (expressed as number of AST nodes generated by the Rhino parser [9]) and the mean translator blowup (i.e., the factor by which the number of AST nodes increases when translating from JavaScript to notJS). The benchmark names are shown in the graphs presented below; the benchmark suite is included in the supplementary material.

The benchmark categories are: **standard**: seven of the large, complex benchmarks from SunSpider [11] and V8 [12] (*mean size*: 2858 nodes; *mean blowup*: $8\times$); **addon**: seven Firefox browser addons, selected from the official Mozilla addon repository [1] (*mean size*: 2597 nodes; *mean blowup*: $6\times$); **generated**: seven programs from the Emscripten LLVM test suite, which translates LLVM bitcode to JavaScript [5] (*mean size*: 38211 nodes; *mean blowup*: $7\times$); and finally **opensrc**: seven real-world JavaScript programs taken from open source JavaScript frameworks and their test suites [3, 8] (*mean size*: 8784 nodes; *mean blowup*: $6.4\times$).

Our goal is to evaluate the precision and performance of JSAI instantiated with several forms of context sensitivity. However, the different sensitivities yield differing sets of function contexts and abstract addresses, making a fair comparison difficult. Therefore, rather than statistical measurements (such as address-set size or closure-set size), we choose a *client-based* precision metric based on an error reporting client. This metric is a proxy for the precision of the analysis.

Our precision metric reports the number of static program locations (i.e., AST nodes) that might throw exceptions, based on the analysis’ ability to precisely track types. JavaScript throws a *TypeError* exception when a program attempts to call a non-function or when a program tries to access, update, or delete a property of **null** or **undef**. JavaScript throws a *RangeError* exception when a program attempts to update the `length` property of an array to contain a value that is not an unsigned 32-bit integer. Fewer errors indicate a more precise analysis.

The performance metric we use is execution time of the analysis. To gather data on execution time, we run each experimental configuration 11 times, discard the first result, then report the median of the remaining 10 trials. We set a time limit of 30 minutes for each run, reporting a timeout if execution time exceeds that threshold. We run all experiments on Amazon Web Services [2] (AWS), using M1 XLarge instances; each experiment is run on an independent AWS instance. These instances have 15GB memory and 8 ECUs, where each ECU is equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor.

We run all 56 analyses on each of the 28 benchmarks, for a total of 1,568 trials (plus the additional 10 executions of each analysis/benchmark pair for the timing data). For reasons of space, we present only highlights of these results. In some cases, we present illustrative examples; the omitted results show similar behavior. In other cases, we deliberately cherry-pick, to highlight contrasts. We are explicit about our approach in each case.

5.2 Observations

For each main sensitivity strategy, we present the data for two trials: the least precise sensitivity in that strategy, and the most precise

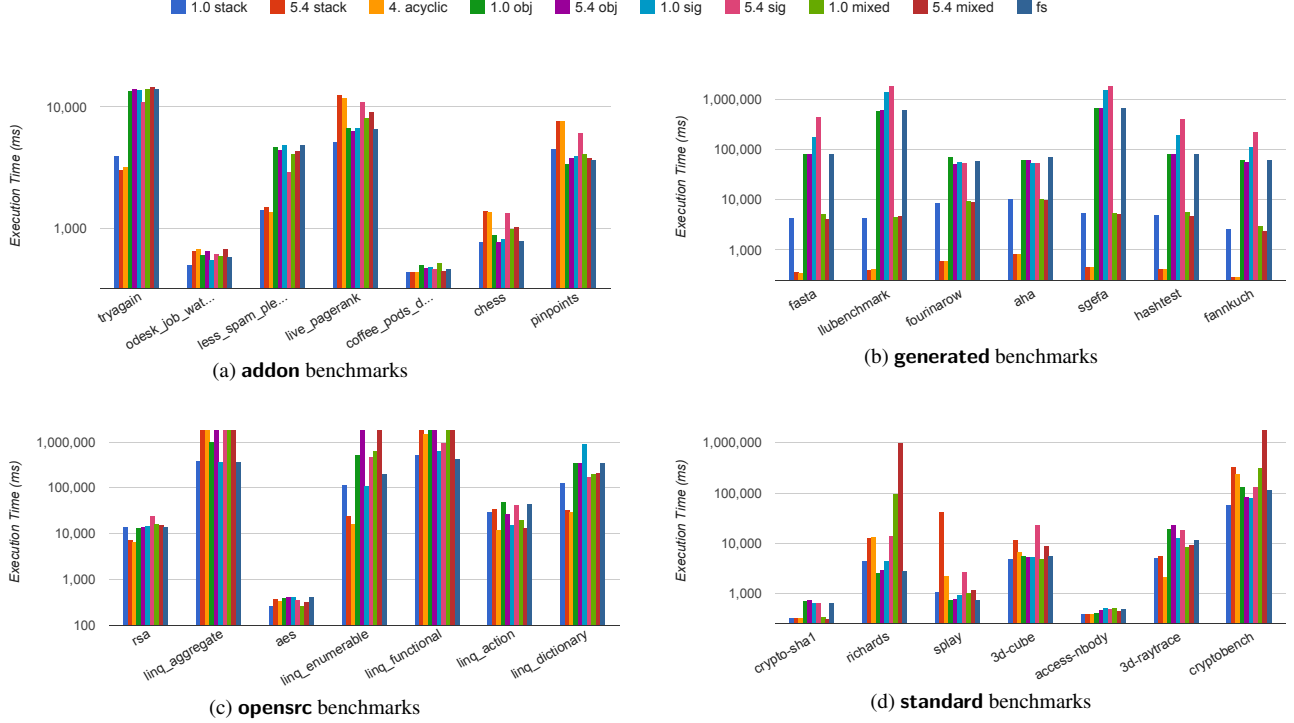


Figure 5: Performance characteristics of different sensitivities across the benchmark categories. The x-axis gives the benchmark names. The y-axis (log scale) gives for each benchmark, the time taken by the analysis (in milliseconds) when run under 10 different sensitivities. Lower bars mean better performance. Timeout (30 minutes) bars are flush with the top of the graph.

sensitivity in that strategy. This set of analyses is: **fs**, **1.0-stack**, **5.4-stack**, **4-acyclic**, **1.0-obj**, **5.4-obj**, **1.0-sig**, **5.4-sig**, **1.0-mixed**, **5.4-mixed**.

Figures 5 and 6 contain performance results, and Figure 7 contains the precision results. The results are partitioned by benchmark category to show the effect of each analysis sensitivity on benchmarks in that category. The performance graphs in Figure 5 plot the median execution time in milliseconds, on a log scale, giving a sense of actual time taken by the various sensitivity strategies. Lower bars are better; timeouts extend above the top of the graph.

We provide an alternate visualization of the performance data through Figure 6 to easily depict how the sensitivities perform relative to each other. Figure 6 is heat map that lays out blocks in two dimensions—rows represent benchmarks and columns represent analyses with different sensitivities. Each block represents relative performance as a color: darker blocks correspond to faster execution time of a sensitivity compared to other sensitivities on the same benchmark. A completely blackened block corresponds to the fastest sensitivity on that benchmark, a whitened block corresponds to a sensitivity that has $\geq 2\times$ slowdown relative to the fastest sensitivity, and the remaining colors evenly correspond to slowdowns in between. Blocks with the red grid pattern indicate a timeout. A visual cue is that columns with darker blocks correspond to better-performing sensitivities, and a row with blocks that have very similar colors indicates a benchmark on which performance is unaffected by varying sensitivities.

Figure 7 provides a similar heat map (with similar visual cues) for visualizing relative precisions of various sensitivity strategies on our benchmarks. The final column in this heat map provides the number of errors reported by the **fs** strategy on a particular benchmark, while the rest of the columns provide the percentage reduction (relative to **fs**) in the number of reported errors due to a corre-

sponding sensitivity strategy. The various blocks (except the ones in the final column) are color coded in addition to providing percentage reduction numbers: darker is better precision (that is, more reduction in number of reported errors). Timeouts are indicated using a red grid pattern.

Breaking the Glass Ceiling. One startling observation is that highly sensitive variants (i.e., sensitivity strategies with high k and h parameters) can be far better than their less-sensitive counterparts, providing improved precision at a much cheaper cost (see Figure 8). For example, on `linq-dictionary`, **5.4-stack** is the most precise *and* most efficient analysis. By contrast, the **3.2-stack** analysis yields the same result at a three-fold increase in cost, while the **1.0-stack** analysis is even more expensive and less precise. We see similar behavior for the `sgafa` benchmark, where **5.4-stack** is an order of magnitude faster than **1.0-stack** and delivers the same results. This behavior violates the common wisdom that values of k and h above 1 or 2 are intractably expensive.

This behavior is certainly not universal,⁵ but it is intriguing. Analysis designers often try to scale up their context-sensitivity (in terms of k and h) linearly, and they stop when it becomes intractable. However, our experiments suggest that pushing past this local barrier may yield much better results.

Callstring vs Object Sensitivity. In general, we find that callstring-based sensitivity (i.e., k,h -**stack** and h -**acyclic**) is more precise than object sensitivity (i.e., k,h -**obj**). This result is unintuitive, since JavaScript heavily relies on objects and object sensitivity was specifically designed for object-oriented languages such as Java. Throughout the benchmarks, the most precise and efficient analyses are the ones that employ stack-based k -CFA. Part of the

⁵For example, `linq-aggregate` times out on all analyses with $k > 1$.

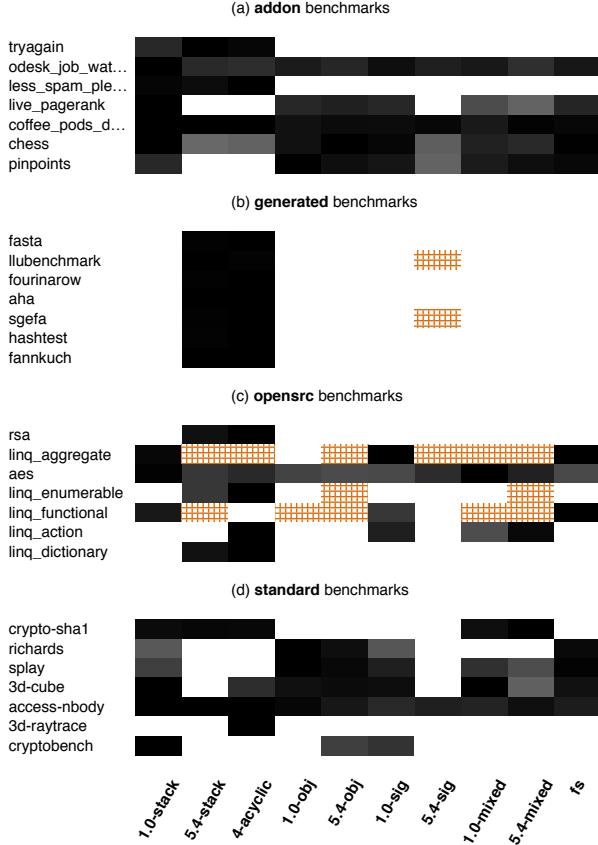


Figure 6: A heat map to showcase the performance characteristics of different sensitivities across the benchmark categories. The above figure is a two-dimensional map of blocks; rows correspond to benchmarks, and columns correspond to analysis run with a particular sensitivity. The color in a block indicates a sensitivities' relative performance on the corresponding benchmark, as compared to fastest sensitivity on that benchmark. Darker colors represent better performance. Completely blackened blocks indicate that the corresponding sensitivity has the fastest analysis time on that benchmark, while completely whitened blocks indicate that the corresponding sensitivity does not time out, but has a relative slowdown of at least $2\times$. The remaining colors are of evenly decreasing contrast from black to white, representing a slowdown between $1\times$ to $2\times$. The red grid pattern on a block indicates a timeout.

reason for this trend is that 25% of the benchmarks are machine-generated JavaScript versions of procedural code, whose structure yields more benefits to callstring-based context-sensitivity. Even among the handwritten open-source benchmarks, however, this trend holds. For example, several forms of callstring sensitivity are more efficient and provide more precise results for the open-source benchmarks than object-sensitivity, which often times out.

Benefits of Context Sensitivity. When it comes to pure precision, we find that more context sensitivity sometimes increases precision and sometimes has no effect. The open-source benchmarks demonstrate quite a bit of variance for the precision metric. A context-sensitive analysis almost always finds fewer errors (i.e., fewer false positives) than a context-insensitive analysis, and increasing the sensitivity in a particular family leads to precision gains. For example, **5.4-stack** gives the most precise error report for `linq_enumerable`, and it is an order of magnitude more precise than a context-insensitive analysis. The addon domain has very little variance for the precision metric, which is perhaps due to shorter

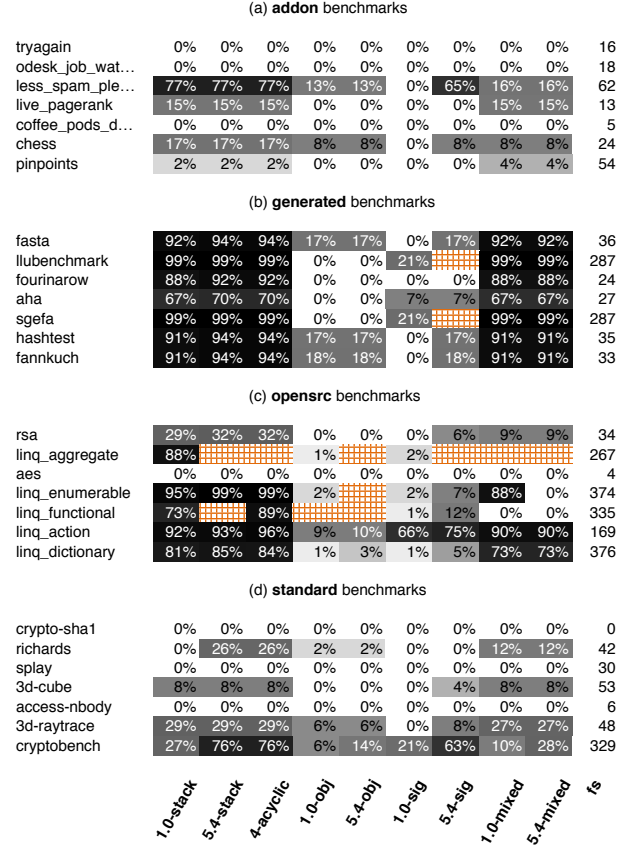


Figure 7: A heat map to showcase the precision characteristics of different sensitivities across the benchmark categories. The above figure is a two-dimensional map of blocks; rows correspond to benchmarks, and columns correspond to analysis run with a particular sensitivity. The rightmost column corresponds to the context insensitive analysis **fs**, and the blocks in this column give the number of errors reported by the analysis under **fs** (which is an upper bound on the number of errors reported across any sensitivity). The color (which ranges evenly from black to white) in the remaining blocks indicate the percentage reduction in number of errors reported by the analysis under the corresponding sensitivity, compared to **fs** on the same benchmark. Darker colors represent more reduction in errors reported, and hence better precision. In addition to the colors, the percentage reduction in errors is also given inside the blocks (higher percentage reduction indicates better precision). The red grid pattern on a block indicates a timeout.

call sequence lengths in this domain. In such domains, it might be wise to focus on performance, rather than increasing precision.

Summary. Perhaps the most sweeping claim we can make from the data is that there is no clear winner across all benchmarks, in terms of JavaScript context-sensitivity. This state of affairs is not a surprise: the application domains for JavaScript are so rich and varied that finding a silver bullet for precision and performance is unlikely. However, it is likely that—within an application domain, e.g., automatically generated JavaScript code—one form of context-sensitivity could emerge a clear winner. The benefit of JSAI is that it is easy to experiment with the context-sensitivity of an analysis. The analysis designer need only implement their base analysis, then instantiate and evaluate multiple instances of the analysis to tune context-sensitivity.

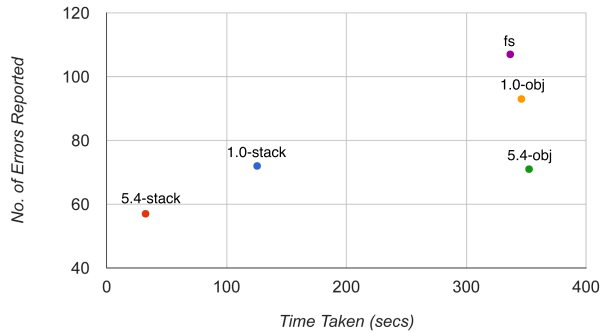


Figure 8: Precision vs. performance of various sensitivities, on the `opensrc linq.dictionaty` benchmark. Interestingly, **5.4-stack** (the most sensitive Stack-CFA analysis) is not only tractable, it exhibits the best performance and the best precision.

5.3 Discussion: JSAI vs. TAJs

Jensen et al.’s Type Analysis for JavaScript [33, 34] (TAJS) stands as the only published static analysis for JavaScript whose intention is to soundly analyze the entire JavaScript language. JSAI has several features that TAJs does not, including configurable sensitivity, a formalized abstract semantics, and novel abstract domains, but TAJs is a valuable contribution that has been put to good use. An interesting question is how JSAI compares to TAJs in terms of precision and performance.

The TAJs implementation (in Java) has matured over a period of five years, it has been heavily optimized, and it is publicly available. Ideally, we could directly compare TAJs to JSAI with respect to precision and performance, but they are dissimilar enough that they are effectively noncomparable. For one, TAJs has known soundness bugs⁶ that can artificially decrease its set of reported type errors. Also, TAJs does not implement some of the APIs required by our benchmark suite, and so it can only run on a subset of the benchmarks. On the flip side, TAJs is more mature than JSAI, it has a more precise implementation of the core JavaScript APIs, and it contains a number of precision and performance optimizations (e.g., the recency heap abstraction [17] and lazy propagation [34]) that JSAI does not currently implement.

Nevertheless, we can perform a qualitative “ballpark” comparison, to demonstrate that JSAI is roughly comparable in terms of precision and performance. For the subset of our benchmarks on which both JSAI and TAJs execute, we catalogue the number of errors that each tool reports and record the time it took for each tool to do so. We find that JSAI analysis time is $0.3\times$ to $1.8\times$ that of TAJs. In terms of precision, JSAI reports from nine fewer type errors to 104 more type errors, compared to TAJs. Many of the extra type errors that JSAI reports are `RangeErrors`, which TAJs does not report due to one of the unsoundness bugs we uncovered. Excluding `RangeErrors`, JSAI reports at most 20 more errors than TAJs in the worst case.

6. Conclusion and Future Work

We have described the design of JSAI, a configurable, sound, and efficient abstract interpreter for JavaScript. JSAI’s design is novel in a number of respects which make it stand out from all previous JavaScript analyzers. We have provided a comprehensive evaluation that demonstrates JSAI’s usefulness. The JSAI implementation and formalisms are freely available as a supplement, and we believe

that JSAI will provide a useful platform for researchers investigating JavaScript analysis.

Our future work includes (1) taking advantage of JSAI’s tunable precision to further investigate what control-flow sensitivities are most useful for JavaScript; (2) writing a number of clients on top of JSAI, including program refactoring, program compression; and (3) extending JSAI to handle language features from the latest ECMA 5 standard.

References

- [1] <https://addons.mozilla.org/en-US/firefox/>.
- [2] <http://aws.amazon.com/>.
- [3] <http://www.defensivejs.com/>.
- [4] <http://doctorjs.org/>.
- [5] <http://www.emscripten.org/>.
- [6] <http://www.drdobbs.com/windows/microsofts-javascript-move/240012790>.
- [7] <http://www.mozilla.org/en-US/firefox/os/>.
- [8] <http://linqjs.codeplex.com/>.
- [9] <https://developer.mozilla.org/en-US/docs/Rhino>.
- [10] <https://github.com/jashkenas/coffee-script/wiki/List-of-languages-that-compile-to-JS>.
- [11] <http://www.webkit.org/perf/sunspider/sunspider.html>.
- [12] <http://v8.googlecode.com/svn/data/benchmarks/v7/run.html>.
- [13] <http://nodejs.org/>.
- [14] <https://developer.mozilla.org/en-US/docs/SpiderMonkey>.
- [15] <http://www.khronos.org/registry/typedarray/spec/latest/>.
- [16] C. Anderson, P. Giannini, and S. Drossopoulou. Towards type inference for javascript. In *European conference on Object-oriented programming*, 2005.
- [17] G. Balakrishnan and T. Reps. Recency-abstraction for heap-allocated storage. In *International conference on Static Analysis*, 2006.
- [18] S. Bandhakavi, N. Tiku, W. Pittman, S. T. King, P. Madhusudan, and M. Winslett. Vetting browser extensions for security vulnerabilities with vex. *Commun. ACM*, 54(9), Sept. 2011.
- [19] R. Chugh, J. A. Meister, R. Jhala, and S. Lerner. Staged information flow for javascript. In *ACM SIGPLAN Conference on Programming Languages Design and Implementation*, 2009.
- [20] R. Chugh, D. Herman, and R. Jhala. Dependent types for javascript. In *International Conference on Object Oriented Programming Systems Languages and Applications*, 2012.
- [21] P. Cousot and R. Cousot. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *ACM Symposium on Principles of programming languages*. ACM Press, New York, NY, 1977.
- [22] P. Cousot and R. Cousot. Systematic design of program analysis frameworks. In *ACM Symposium on Principles of Programming Languages*, 1979.
- [23] ECMA. *ECMA-262: ECMAScript Language Specification*. Third edition, Dec. 1999. URL <http://www.ecma.ch/ecma1/STAND/ECMA-262.HTM>.
- [24] P. A. Gardner, S. Maffeis, and G. D. Smith. Towards a program logic for javascript. In *ACM Symposium on Principles of programming languages*, 2012.
- [25] S. Guarnieri and B. Livshits. Gatekeeper: mostly static enforcement of security and reliability policies for javascript code. In *Conference on USENIX security symposium*, 2009.
- [26] A. Guha, S. Krishnamurthi, and T. Jim. Using static analysis for Ajax intrusion detection. In *World Wide Web Conference*, 2009.

⁶ We uncovered several soundness bugs when we were formalizing our semantics, and the TAJs authors confirmed them as errors.

- [27] A. Guha, C. Saftoiu, and S. Krishnamurthi. The essence of javascript. In *European conference on Object-oriented programming*, 2010.
- [28] A. Guha, C. Saftoiu, and S. Krishnamurthi. Typing local control and state using flow analysis. In *European conference on Programming languages and systems*, 2011.
- [29] B. Hardekopf, B. Wiedermann, B. Churchill, and V. Kashyap. Widening for control-flow. In *International Conference on Verification, Model Checking, and Abstract Interpretation*, 2014.
- [30] P. Heidegger and P. Thiemann. Recency types for analyzing scripting languages. *European conference on Object-oriented programming*, 2010.
- [31] K. Ishizaki, M. Kawahito, T. Yasue, H. Komatsu, and T. Nakatani. A study of devirtualization techniques for a java just-in-time compiler. In *ACM International Conference on Object Oriented Programming Systems Languages and Applications*, 2000.
- [32] D. Jang and K.-M. Choe. Points-to analysis for javascript. In *Symposium on Applied Computing*, 2009.
- [33] S. H. Jensen, A. Møller, and P. Thiemann. Type Analysis for Javascript. In *International Symposium on Static Analysis*, 2009.
- [34] S. H. Jensen, A. Møller, and P. Thiemann. Interprocedural Analysis with Lazy Propagation. In *International Symposium on Static Analysis*, 2010.
- [35] S. H. Jensen, P. A. Jonsson, and A. Møller. Remedying the Eval that Men Do. In *International Symposium on Software Testing and Analysis*, 2012.
- [36] V. Kashyap and B. Hardekopf. Security signature inference for javascript-based browser addons. In *Symposium on Code Generation and Optimization*, 2014.
- [37] V. Kashyap, J. Sarracino, J. Wagner, B. Wiedermann, and B. Hardekopf. Type refinement for static analysis of javascript. In *Symposium on Dynamic Languages*, 2013.
- [38] H. Lee, S. Won, J. Jin, J. Cho, and S. Ryu. Safe: Formal specification and implementation of a scalable analysis framework for ecma-script. In *International Workshop on Foundations of Object-Oriented Languages*, 2012.
- [39] F. Logozzo and H. Venter. Rata: Rapid Atomic Type Analysis by Abstract Interpretation – Application to Javascript Optimization. In *Joint European Conference on Theory and Practice of Software, International Conference on Compiler Construction*, 2010.
- [40] M. Madsen, B. Livshits, and M. Fanning. Practical static analysis of JavaScript applications in the presence of frameworks and libraries. In *ACM Symposium on the Foundations of Software Engineering*, Aug. 2013.
- [41] S. Maffei, J. C. Mitchell, and A. Taly. An operational semantics for javascript. In *Asian Symposium on Programming Languages and Systems*, 2008.
- [42] F. Meawad, G. Richards, F. Morandat, and J. Vitek. Eval begone!: semi-automated removal of eval from javascript programs. In *ACM International Conference on Object Oriented Programming Systems Languages and Applications*, 2012.
- [43] M. Schäfer, M. Sridharan, J. Dolby, and F. Tip. Dynamic determinacy analysis. In *ACM SIGPLAN Conference on Programming Languages Design and Implementation*. ACM, 2013.
- [44] I. Sergey, D. Devriese, M. Might, J. Midtgaard, D. Darais, D. Clarke, and F. Piessens. Monadic abstract interpreters. In *ACM SIGPLAN Conference on Programming Languages Design and Implementation*. ACM, 2013.
- [45] Y. Smaragdakis, M. Bravenboer, and O. Lhoták. Pick your contexts well: understanding object-sensitivity. In *ACM Symposium on Principles of programming languages*, 2011.
- [46] A. Taly, U. Erlingsson, J. C. Mitchell, M. S. Miller, and J. Nagra. Automated analysis of security-critical javascript apis. In *IEEE Symposium on Security and Privacy*, 2011.
- [47] P. Thiemann. Towards a Type System for Analyzing Javascript Programs. In *European Conference on Programming Languages and Systems*, 2005.
- [48] D. Van Horn and M. Might. Abstracting abstract machines. In *International Conference on Functional Programming*, 2010.